Pizza Toss. *World pizza champion Tony Gemigniani demonstrates his pizza-tossing form. In a pizza toss, the angular momentum vector is very close to the normal vector. Thus the pizza never turns over, and always lands "heads". (Photo courtesy of Tony F. Gemigniani, President, World Pizza Champions, Inc.)*

The Fifty-one Percent Solution

FOR CENTURIES, COINS HAVE BEEN AN ICON of randomness. Who hasn't flipped a coin to decide between two equally appealing alternatives—which restaurant to go to, which road to take? Especially when the choice doesn't matter too much, tossing a coin beats thinking.

But sometimes, even important choices are left to the caprice of a coin. In 1845, two settlers in the Oregon Territory, Asa Lovejoy and Francis Pettygrove, founded a new town on the banks of the Willamette River. Lovejoy wanted to name it after his birthplace, Boston, but Pettygrove preferred to name it after his own hometown—Portland, Maine. A coin toss seemed like the fairest way to settle the dispute. The penny landed in Pettygrove's favor, and that is why the largest city in Oregon is now called Portland, instead of Boston.

The implicit assumption behind coin flips has always been that that they are fair—in other words, that heads and tails have an equal chance of occurring. However, a team of three mathematicians has now proved that this assumption is incorrect. Any coin that is tossed vigorously and high, and caught in midair (rather than bouncing on the ground) has about a 51 percent chance of landing with the same face up that it started with. Thus, if you catch a glimpse of the coin before it is

Persi Diaconis and Susan Holmes.

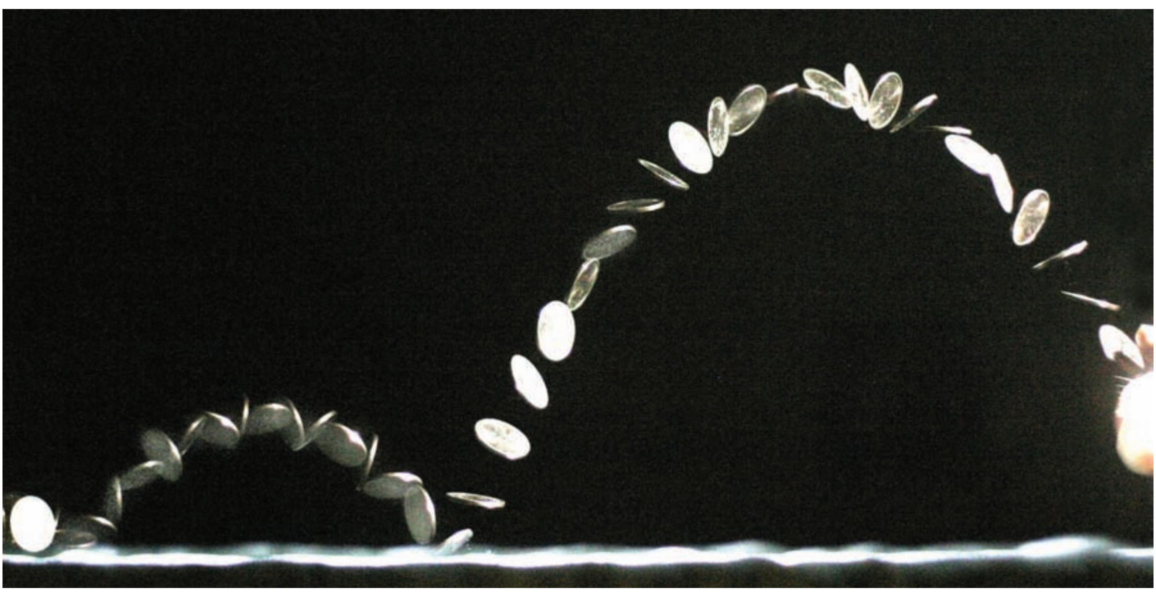

Figure 1. *A stroboscopic photo of a coin toss over a period of one second. Between the toss and the first landing, the coin made two full revolutions (or four half-revolutions), and thus the upward face was alternately heads-tails-heads-tails-heads. Thus, it landed in the same orientation that it started, a result that, according to new research, happens about 51 percent of the time. (Photo courtesy of Andrew Davidhazy, Rochester Institute of Technology.)*

tossed and see heads up, you should call heads; on the other hand, if you see the tails side up, you should call tails. This will give you a 51 percent chance of predicting the outcome correctly.

Persi Diaconis and Susan Holmes of Stanford University, together with Richard Montgomery of the University of California at Santa Cruz, published their discovery in *SIAM Review* in 2007, although it was originally announced in 2004. According to Diaconis, who has given public lectures on the topic numerous times, the result is a difficult one for most non-mathematicians to grasp. The coin's bias has nothing to do with whether it is weighted more on one side or the other, nor with any asymmetry in its shape. In fact, their analysis assumes that the coin is perfectly symmetrical. The coin's bias lies not in its shape, but in its motion—the dynamics of a rigid, rotating object.

From a physicist's point of view, a coin's motion is completely deterministic. If you know the initial orientation, velocity, and angular velocity of the coin, you can predict its future flight perfectly. Diaconis, in fact, has built a coin-tossing machine that illustrates this fact (see Figure 2). The machine tosses heads with spooky consistency. You press a lever, it launches a coin into a collecting cup, and the coin will come up heads every single time. There's no chance involved.

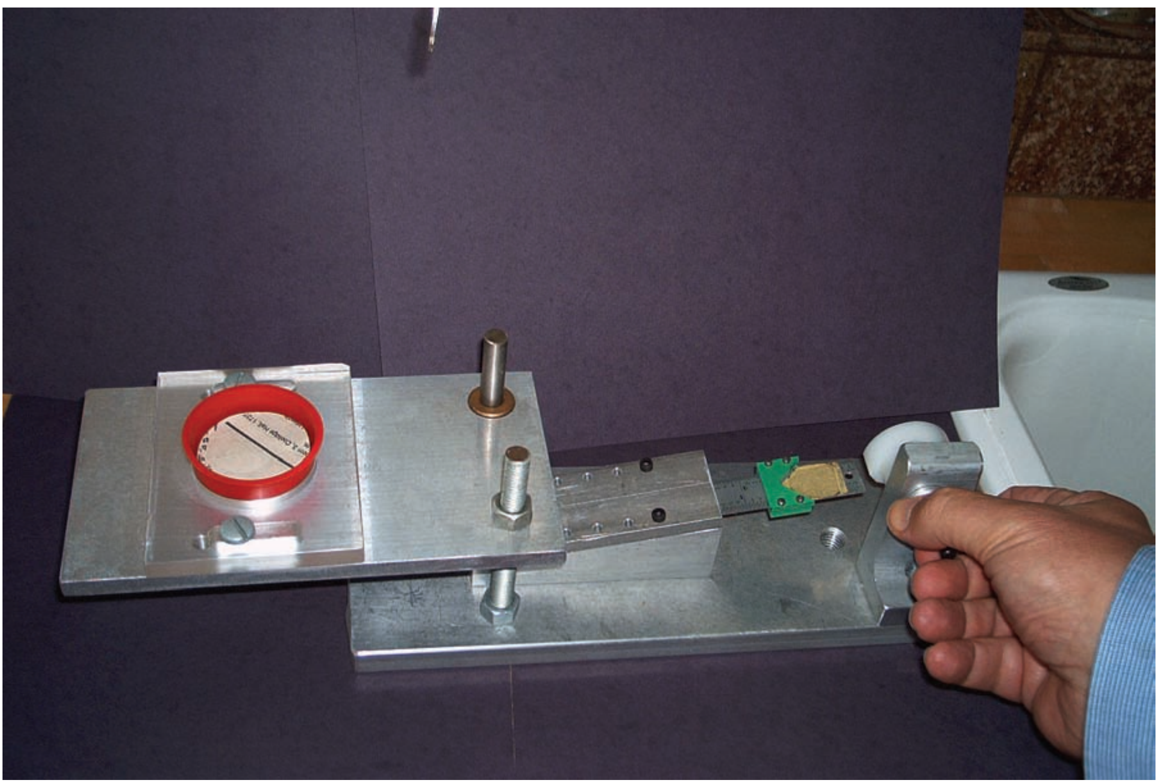

Figure 2. *Persi Diaconis and his colleagues have built a coin tosser that throws heads 100 percent of the time. A coin's flight is perfectly deterministic—it is only our lack of machine-like motor control that makes it appear random. (Photo courtesy of Susan Holmes.)*

If a machine can produce heads 100 percent of the time, why do humans have such blind faith in the coin's randomness? In a sense, it is not the coin's randomness that is at issue, but our own clumsiness. To produce heads all the time, you need extremely precise control over the coin's initial conditions, such as the strength of the toss and the rotation rate imparted to the coin. Humans do not normally have such fine motor control. Even so, there are some ways that humans can, either intentionally or unintentionally, bias their coin tosses.

First, if you toss the coin in the same way that a pizza maker tosses his pizza dough (see **Pizza Toss**, page 34)—setting it in motion around its normal axis instead of its diameter—then the coin will never flip over, and you will get heads 100 percent of the time (assuming the coin starts with heads up). Magicians have perfected a trick based on the "pizza toss." They realized that they could add a little wobble, but not too much. The casual viewer cannot tell the difference between the wobbly "pizza toss" and a real coin flip, and the magician will still get heads 100 percent of the time.

A second form of bias occurs if you give the coin a very wimpy toss, so that it rotates only a half-turn before landing. The coin will always land with the opposite face pointing up, and so a sequence of "wimpy tosses" will go heads-tails-heads-tails forever. Even though the percentage of heads is exactly 50 percent, these tosses are far from random.

Diaconis once received a class project, called "The Search for Randomness," that fell into this second trap. A teacher had asked his students to flip a coin 300 times each, thus generating a table of 10,000 supposedly random coin flips, which he sent to Diaconis. Alas, Diaconis says, "The results were very patterned. The reason was that the students got bored." (Wouldn't you, if you were flipping a coin for a full class period?) The more bored the students got, the wimpier their flips were, and the more frequently the telltale pattern heads-tails-heads-tails started to show up.

Both of these exceptions turn out to be very important for the mathematics of coin-flipping because they are extreme cases. The first exception shows that it makes a difference what axis the coin rotates about. The second shows the importance of making sure the coin spins a reasonably large number of times before it is caught. By trickery or by simple disinterest, a human can easily manage a biased coin flip. But the question Diaconis wanted to know was: What happens if an ordinary human honestly tries to achieve a random and unbiased flip? Can he do it?

Clearly, the flip should rotate the coin about its diameter, and it should be vigorous enough to ensure a fairly large number of revolutions. Under those two conditions, Joe Keller, an applied mathematician at Stanford, showed in 1986 that the probability of heads does indeed approach 50 percent. However, this was a somewhat idealized result because Keller assumed that the number of revolutions approaches infinity. Diaconis' student Eduardo Engel followed up on Keller's work by studying what would happen under more realistic conditions—a coin that rotates between 36 and 40 times per second, for about half a second (to be precise, between .44 and .56 seconds). Engel, who is now at Yale University, showed that the

If a machine can produce heads 100 percent of the time, why do humans have such blind faith in the coin's randomness? In a sense, it is not the coin's randomness that is at issue, but our own clumsiness.

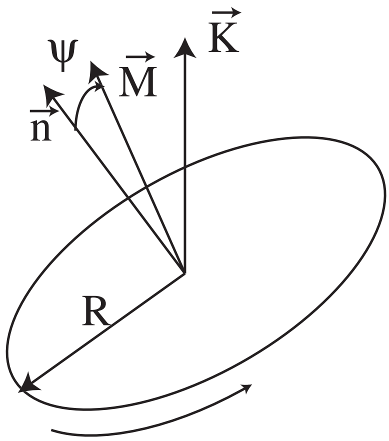

Figure 3. *The three relevant vectors for determining how a coin will land are the normal vector (n), the angular momentum vector M, and the upward vector K. The vector M remains stationary, and n precesses around it in a cone. Thus, the angle between M and n remains constant (ψ). The coin will turn over only if the angle between n and K exceeds 90 degrees at some time. In this sketch, because ψ is small, the coin will never turn over—this is a "pizza toss". (Figure courtesy of Susan Holmes.)*

probability of heads under these circumstances (assuming the coin starts with heads up) lies between 0.444 and 0.556. Engel's result was already quite sophisticated, requiring careful estimates of integrals over the space of allowable initial conditions. Neither he nor Keller discovered any indication that coin tosses are biased.

However, there is a second simplifying assumption in both Keller's and Engel's work. They assumed that the flip sets the coin spinning *exactly* around its diameter (i.e., the axis of rotation lies exactly in the plane of the coin). Even if the coin-flipper is trying to give the coin an honest flip, it is unrealistic to expect such perfection. What happens if the coin rotates about an axis that is neither perpendicular to the coin as in the "pizza toss," nor exactly in the plane of the coin as in the "Keller toss," but instead points off in some oblique direction? That is exactly when the problem gets interesting.

In 2003, Diaconis visited Montgomery in Santa Cruz, and saw on the wall of his office a poster depicting how a falling cat rotates itself to land on its feet. Suddenly he knew he had found the right person to ask about the rotation of a falling coin.

Coins are actually simpler than cats because they are rigid objects, and also because they have circular symmetry. Cats, according to Montgomery, are best viewed as a system of two rigid objects (the cat's front and rear) with a flexible link between them. The flexible link—the cat's muscles—gives it control over its landing orientation.

The coin, on the other hand, has no such control. Its motion is completely determined from the moment it is tossed to the moment it is caught. The motion of a "free rigid object," as physicists call it, has been understood ever since Leonhard Euler in the eighteenth century. Although, the description of the motion is quite simple, it is perhaps not as familiar as it should be.

From the point of view of a physicist, the key parameter for describing the coin's motion is its angular momentum—or what we have loosely referred to as the "axis of rotation" above. To the untrained observer, the coin's behavior looks like a complicated combination of spinning, tumbling, and wobbling, but to a physicist, all of these kinds of motion can be summed up in one vector, the angular momentum. Moreover, the angular momentum vector remains unchanged for the entire time that the coin is in the air.

What matters most for the outcome of the coin toss, though, is not the orientation of the angular momentum, but the orientation of the coin's normal vector. (This is the unit vector pointing perpendicularly to the coin, in the "heads" direction, as shown in Figure 3.) If the coin is caught with the normal vector pointing anywhere in the upper hemisphere, it will be interpreted as a "heads" flip. If the normal vector points towards the lower hemisphere, the flip will be recorded as "tails."

The normal vector and the angular momentum vector are different, but they are related in an exceptionally simple way. The angle between them, denoted ψ, remains constant. This means that the normal vector *precesses* around the fixed angular momentum vector, forming a cone with an opening angle of ψ radians. Alternatively, when the trajectory of the unit normal vector is plotted on a sphere, the result is a circle of radius ψ

whose center is the angular momentum vector (normalized to have length one).

From this point of view, the "pizza flip" and the "Keller flip" are particularly simple cases. In the "pizza flip," the angular momentum vector and the normal vector coincide. The angle between them is 0. Because the angular momentum is constant, and the normal vector moves in a "circle of radius 0" around it, the normal vector is also constant. If the coin starts out heads (i.e., with normal vector $(0, 0, 1)$) then the normal vector will continue to point in the same direction, and the coin will always land heads.

In the "Keller flip," on the other hand, the angular momentum vector lies in the plane of the coin, making an angle of 90° or $\pi/2$ radians with the normal vector. Hence the normal vector precesses around a circle of radius $\pi/2$—in other words, a great circle on the unit sphere. If we assume that it starts pointing in the direction $(1, 0, 0)$, then it will rotate at a constant rate from the north pole to the south pole and back. Half of the time the normal vector will be in the upper hemisphere, and half of the time it will be in the lower hemisphere. In a rough sense, this explains why the coin has a 50 percent chance of landing heads and a 50 percent chance of landing tails (as Keller proved).

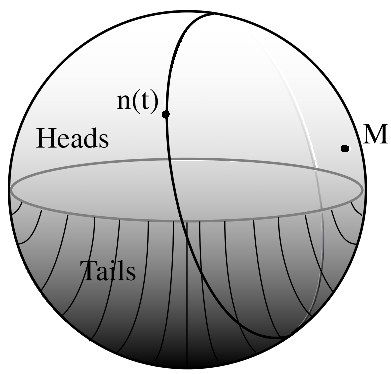

Figure 4. *In general, the normal vector will describe a circle centered at M while the coin is in the air. If the coin starts perfectly horizontal (so that n is perfectly vertical), then n will inevitably spend more time in the upper hemisphere than the lower hemisphere. If the coin is caught at a random time, it will therefore have a greater than 50 percent probability of being caught with n in the upper hemisphere. (Figure courtesy of Susan Holmes.)*

In a real flip, the normal vector will precess in a circle of radius ψ radians, a "small circle" rather than a great circle (See Figure 4). If the coin is tossed from a heads-up position, then this small circle passes through the point $(1, 0, 0)$. If ψ is less than $\pi/4$ radians, then in fact the entire small circle lies in the upper hemisphere. This means that the coin never actually turns over; it will land heads with 100% probability. This is why magicians do not have to execute a perfect "pizza toss" in order to ensure the coin lands heads; there is in fact considerable room for error. A coin tossed with ψ close to $\pi/4$ will precess so vigorously that the eye is easily deceived into thinking that it is tumbling, and yet it will always land heads.

As beautiful as it is, this analysis did not allow the mathematicians to make a precise estimate of the amount of bias in a human-tossed coin . . . To obtain a quantitative estimate, therefore, the mathematicians had to study empirically how humans actually flip coins. And that was not nearly as easy as it sounds.

If ψ lies between $\pi/4$ and $\pi/2$, then the normal vector will sometimes enter the lower hemisphere. Thus there will be a nonzero chance of the coin landing tails. However, in all cases the normal vector will spend more time in the upper hemisphere than in the lower hemisphere, and therefore the probability of heads will always exceed 50 percent. Although this is a somewhat informal argument, Diaconis, Holmes and Montgomery formalized it in the same way that Keller did, by taking a limit as the time of flight goes to infinity. The probability of heads (assuming that the coin starts with heads up) is

$$p(\psi, \phi) = \frac{1}{2} + \frac{1}{\pi} \sin^{-1}(\cot \psi \cot \phi),$$

where ϕ is the (fixed) angle that the angular momentum vector makes with the vertical. If the inverse sine is undefined, as is the case when ψ is small, then the probability of heads is simply 1.

As beautiful as it is, this analysis did not allow the mathematicians to make a precise estimate of the amount of bias in a human-tossed coin. That depends on how close ψ is to $\pi/2$ (i.e., how closely the coin flipper can approximate a perfect "Keller flip"). To obtain a quantitative estimate, therefore, the mathematicians had to study empirically how humans actually flip coins. And that was not nearly as easy as it sounds.

"When Joe [Keller] had written his paper, I wanted an answer to the question of how many times the coin turns," Diaconis says. "At that time, I found out that there was only one slow-motion camera at Stanford, which was owned by the football team. In order to use it, I would have to pay the operator something like $2800 for a two-hour session. I wanted to know the answer, but I didn't want to know it that badly!"

A decade and a half later, after they had talked with Montgomery, Diaconis and Holmes again tried to videotape real coin flips. But an ordinary video camera is far too slow: it shoots 60 frames per second. Because a typical coin makes at least 20 full revolutions per second, a frame-by-frame data set is far too coarse to tell what it is actually doing. "You'd like to have up to 800 frames per second," Diaconis says. They tried to overcome the problem by using eight videocameras at once, but the logistics proved to be too difficult. "I was in despair," Diaconis says. But this time, he found out that Abbas El Gamal, of Stanford's electrical engineering department, had an ultra-slow-motion camera. Unlike the football team, El Gamal was happy to have his camera used for coin-flipping research. Problem solved!

In reality, that was only the beginning. From the physical point of view, it wasn't so easy to flip a coin in such a way that the camera could record its flight successfully. The camera was stationary, so the flip had to pass directly in front of the camera lens. And the camera would only record for about a quarter of a second, so the flip had to be synchronized very closely to the start of filming. Out of 50 attempts, only 27 gave useful results.

Then there were mathematical challenges, some with very ingenious solutions. First was the problem of simply finding the coin in each of the images. With 27 videos of 100 frames each, it would be impractical to draw the outline of the coin on each frame by hand. Moreover, the human eye and hand are not objective enough. Holmes tried out several different statistical learning algorithms to find one that would be able to pick out the coin pixels from the background pixels; she then used standard statistical techniques to find the ellipse that most closely matched the edge of the coin. Finally, she could use this ellipse to determine which way the normal vector was pointing at each instant.

At this point, a surprise awaited her. When plotted on a sphere, the normal vectors did not lie along a circle (the circle of precession), as theory said they were supposed to. Instead, they just made a formless blob. (See Figure 5.) “It was devastating the first time we saw the data,” Holmes recalls.

Finally, Holmes figured out what the problem was. The two-dimensional photographic image of the coin does not completely determine its orientation in space. For any given ellipse, there are four possible orientations of the coin that would produce that ellipse: the nearer edge of the coin could be either above or below the farther side, and the top face could either be heads or tails. If you know the coin’s history, you can tell which orientation is correct because the coin won’t suddenly jump from one orientation to another. But teaching the computer to make this determination automatically took a little bit of work. Finally Holmes was able to unscramble the data enough to get a satisfactory approximation of a circle. Once the circle had been determined, it was easy to find the angular momentum vector (the center of the circle) and the parameter ψ (the angular radius of the circle).

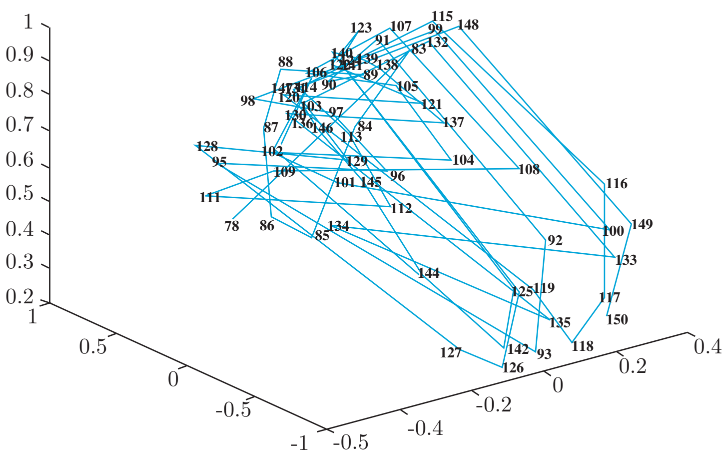

Figure 5. *At first, the photographic data seemed to show the normal vectors traveling in a strange, jerky pattern—not in a circle as Figure 4 would predict. However, the researchers eventually figured out that the problem lay in the ambiguity of a 2-dimensional photographic image. The normal to the coin could be pointing in any of four directions in 3-space, and their automated image analysis program was picking the wrong ones. (Figure courtesy of Susan Holmes.)*

As it turned out, for some of the tosses there was an independent way to check the computation of ψ. It’s an intriguing, and not obvious, physical fact that each time the coin’s normal vector precesses exactly one time about the angular momentum

The average probability was 0.508, which they rounded up to 0.51, and this was the basis for their claim that real coins have a 51 percent chance of landing with the same side up that they began with.

vector, the coin also rotates a fixed amount (in a coin-centered coordinate frame) about the normal vector. Figure 6 shows a nice example: the coin was imaged three times facing in the same direction. (The images were separated by 20 frames, and therefore by 1/30 of a second). Between each pair of frames, the letter "T" in the center of the coin has rotated through an equal angle. This angle of rotation, ΔA, is related to ψ by the following equation:

$$\Delta A = -(1 - I_1/I_3)2\pi \cos(\psi),$$

where I_1 and I_3 represent the moments of inertia of the coin about its diameter and normal, respectively. For a very thin coin, the ratio I_1/I_3 is $\frac{1}{2}$, but for a U.S. half dollar, the ratio is about 0.513.

The effect is well known to quantum physicists, who call it the *Berry phase* after Michael Berry, who (re-)discovered it in 1984. It can arise whenever the wave parameters of an oscillating system are changed slowly and then returned to their original values. The system will appear to have returned to its original state, but it will have accumulated a phase difference that can be detected (in the quantum physics applications) by using an interferometer. Curiously, Montgomery had just written a book with a chapter devoted to Berry phases—another way in which he turned out to be just the right man for the project.

The Berry phase gave Diaconis, Holmes, and Montgomery an independent way of computing the angle ψ, and thereby validating that their image analysis algorithm was working correctly. With their computation of ψ and the angular momentum vector, they could use Diaconis' formula for $p(\psi, \phi)$ to compute the probability of heads on each of the 27 flips that they had recorded. (Note that the actual outcome of the flips was immaterial—it was the probability that they wanted.) The average probability was 0.508, which they rounded up to 0.51, and this was the basis for their claim that real coins have a 51 percent chance of landing with the same side up that they began with.

The result comes with a number of caveats, none of which alter the basic conclusion that coin tosses are biased. First, it is exceedingly important to catch the coin in midair, and not let it hit the ground. Once it hits the ground, other factors—such as the shape or weight distribution of the coin—start playing a role. Magicians have learned how to shave the edge of a coin so that if you spin it on a table (rather than tossing it in the air) the coin will always come up heads. Although letting the coin bounce may be acceptable for a perfectly symmetric coin, it is too easy to tamper with the coin and change the probabilities.

Also, the 51 percent probability is only an ideal estimate, when the coin is allowed to fly for a long time. Real coin flips tend to last only a half second or so. The finite-time-of-flight effect means that the actual probability of heads falls in a range of values (just as Engel's analysis of Keller flips gave the probability a range of values). However, the range will be centered on 51 percent, not on 50 percent.

Third—and this is the point that worries Diaconis the most—it's still unclear just how representative their 27 tosses were of what happens in a real coin toss. Because it was so hard to synchronize the tosses with the high-tech camera, the tosses were probably performed more carefully and perhaps less vigorously than real coin tosses. Thus it is quite possible that the real bias in favor of the starting position is more than 51 percent, and indeed quite a bit more. This became apparent in a "low-tech" experiment that Diaconis performed, attaching a ribbon to the coin so that he could count how many times it flipped over. (The number of flips is the same as the number of twists in the ribbon when the coin is caught.) In 4 out of 100 tosses, the coin never flipped at all—Diaconis had unintentionally performed a "pizza flip." In 3 out of 100 tosses, the coin flipped only once. Thus a significant number of real-world coin flips may be executed unintentionally in a way that makes them far from random.

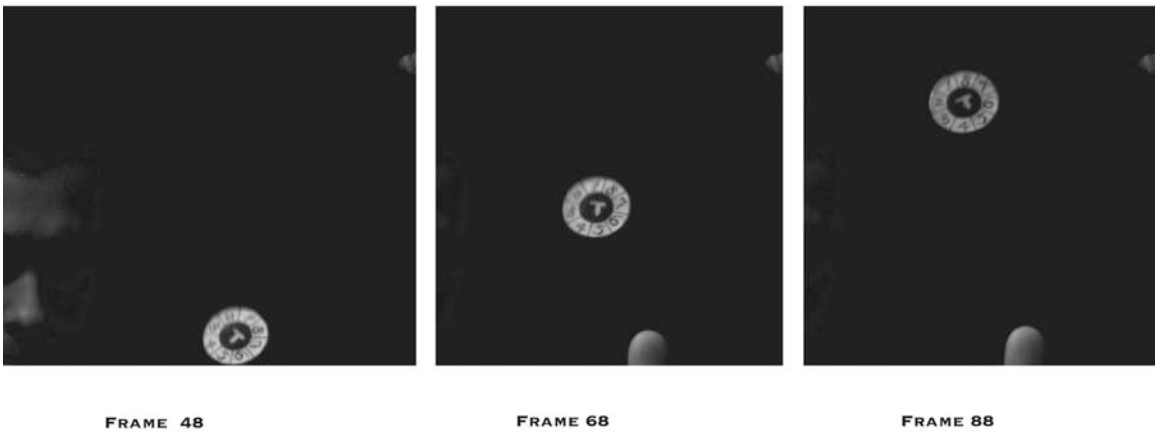

Figure 6. *In this trial, the tosser happened to flip the coin at very close to 30 revolutions per second. Thus, with a high-speed camera that photographed the coin 600 times per second, the coin was nearly in the same orientation every 20 frames. Notice that while the coin has made a complete revolution, it has precessed around its normal axis by less than a full revolution. In fact, the amount of precession, called the Berry phase, provides an independent way to measure the angle ψ mentioned in Figure 3. (Figure courtesy of Susan Holmes.)*

Montgomery's wife Judith, a school math teacher, has suggested holding a "great California flip-off" to see if real-world coin flips actually do have the predicted amount of bias. It will take a lot of flips, though. To verify experimentally that the probability of heads is *not* 50 percent but 51 percent, one would need a random sample of about a quarter-million flips. Given the difficulty of even obtaining a sample of 10,000 flips (remember the bored math students?), Diaconis is not optimistic about the chances of pulling off such a massive experiment. "The idea of performing quality control on that data puts me off," he says.

For all three researchers, the coin-flipping work has led in unexpected directions, some of them much more serious than the original problem. To Montgomery, it's a great teaching example for a course in differential geometry. The rotations of the coin correspond to elements of the Lie group SO(3), which can be associated to points on a hypersphere (that is, a

Inspiration for big things can indeed come from humble sources, even the common flip of a coin or a dinner plate.

three-dimensional sphere in four-dimensional space). Different shapes or weightings of the coin correspond to different ways of measuring distance on the hypersphere, and the complicated dance of the coin always represents the shortest path through this three-dimensional space.

For Diaconis, a sequence of coin flips is an analogue for a much more complicated process—namely, the folding of a protein. The coin exhibits a very simple kind of dependence between its successive states—namely, it has a 51 percent chance of staying in the same state it was in (heads or tails), and a 49 percent chance that it will switch to the opposite state. Yet determining this dependence from first principles was not at all easy. A protein molecule has many more configurations or states; it's not just a simple dichotomy of "heads" and "tails." The challenge is to simplify the description of the molecule so that the number of states is manageable, the states are still physically meaningful, and one can compute the probability of moving from one state to another. "The coins were a direct motivation for our work on protein folding," Diaconis says.

Finally, for Holmes, the "baby problem" of automatically detecting a coin in a photograph has led to a new interest in image analysis. She is currently working with a software package called Gemident, to train computers to recognize cancer and immune cells in an image of a lymph node (Figure 7). "In the beginning, everyone counted these cells by hand, and it wasn't an objective method," she says. The automated detection algorithm makes it possible to do quantitative statistical analysis of the images—for example, determining where the cancer cells lie in relation to the immune cells (such as T-cells and B-cells).

"The value [of the coin-flipping project] for me was that it showed me it was doable, that you can work with immense computer files like these," Holmes says. "Now I'm able to say to the biologists, 'You shouldn't do this by hand, because I can teach the computer to do it."'

Perhaps they will follow in the footsteps of Richard Feynman, the physicist who first worked out the theory of quantum electrodynamics—and who said that his theory was inspired by watching a dinner plate that was tossed into the air. In *Surely You're Joking, Mr. Feynman*, the Nobel laureate recalls, "The [Feynman] diagrams and the whole business that I got the Nobel Prize for came from that piddling around with the wobbling plate." Inspiration for big things can indeed come from humble sources, even the common flip of a coin or a dinner plate.

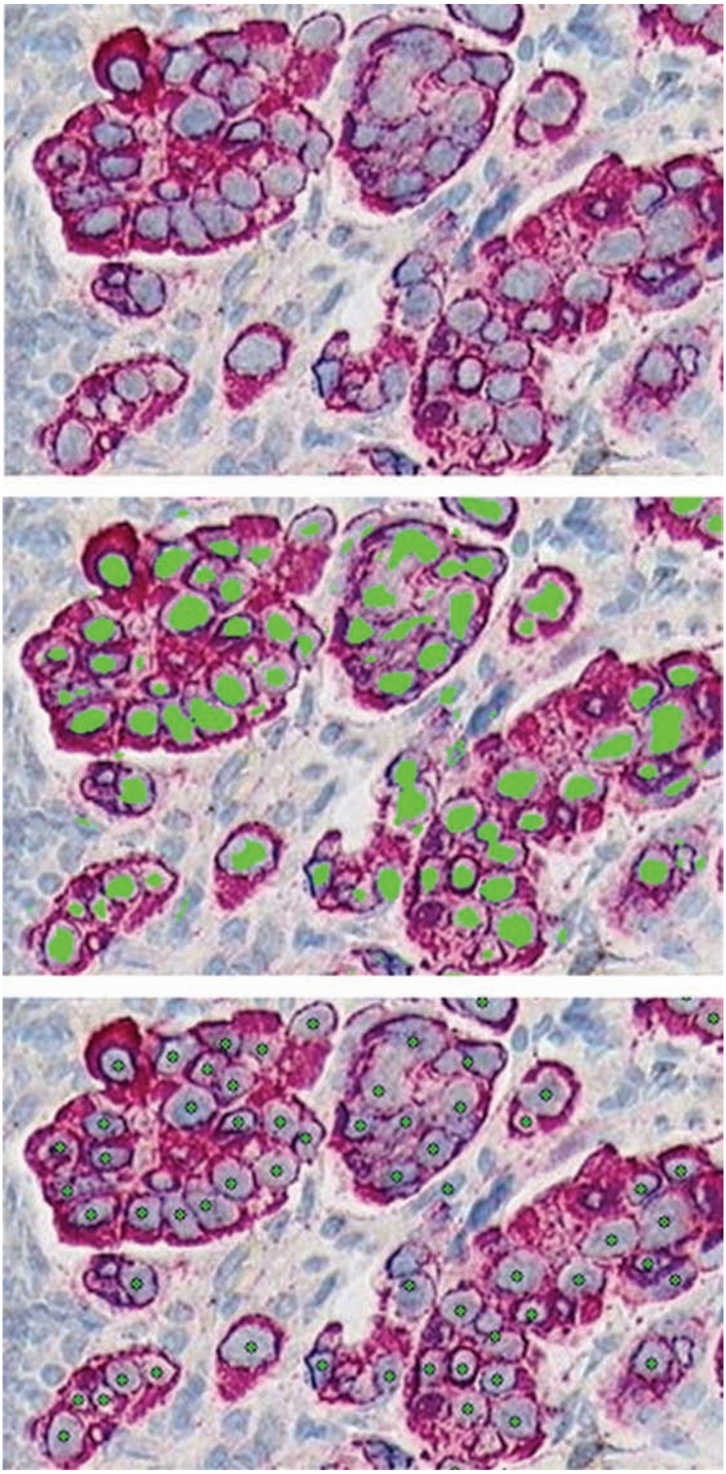

Figure 7. *Since the group's coin experiments, Holmes has worked on other projects involving automated image analysis. Here, she has "taught" a computer to recognize cancerous cells (indicated by green dots) in a microscopic image of a lymph node. (Figure courtesy of Susan Holmes.)*